SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: Skatrud, Paul L.
                 Peery, Robert B.
                 de Waard, Maarten

    (ii) TITLE OF INVENTION: Multiple Drug Resistance Gene atrD of
           Aspergillus Nidulans

    (iii) NUMBER OF SEQUENCES: 3

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: Eli Lilly and Company
        (B) STREET: Lilly Corporate Center
        (C) CITY: Indianapolis
        (D) STATE: Indiana
        (E) COUNTRY: U.S.
        (F) ZIP: 46285

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER:
        (B) FILING DATE:
        (C) CLASSIFICATION:

    (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: Webster, Thomas D.
        (B) REGISTRATION NUMBER: 39,872
        (C) REFERENCE/DOCKET NUMBER: X-11766

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: 317-276-3334
        (B) TELEFAX: 317-276-2763

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4002 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: NO

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION: 1..4002

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

17

```
    ATG TCC CCG CTA GAG ACA AAT CCC CTT TCG CCA GAG ACT GCT ATG CGC        48
    Met Ser Pro Leu Glu Thr Asn Pro Leu Ser Pro Glu Thr Ala Met Arg
     1               5                  10                  15

 5  GAA CCT GCT GAG ACT TCA ACG ACG GAG GAG CAA GCT TCT ACA CCA CAC        96
    Glu Pro Ala Glu Thr Ser Thr Thr Glu Glu Gln Ala Ser Thr Pro His
                    20                  25                  30

    GCT GCG GAC GAG AAG AAA ATC CTC AGC GAC CTC TCG GCT CCA TCT AGT       144
10  Ala Ala Asp Glu Lys Lys Ile Leu Ser Asp Leu Ser Ala Pro Ser Ser
                35                  40                  45

    ACT ACA GCA ACC CCC GCA GAC AAG GAG CAC CGT CCT AAA TCG TCG TCC       192
15  Thr Thr Ala Thr Pro Ala Asp Lys Glu His Arg Pro Lys Ser Ser Ser
        50                  55                  60

    AGC AAT AAT GCG GTC TCG GTC AAC GAA GTC GAT GCG CTT ATT GCG CAC       240
    Ser Asn Asn Ala Val Ser Val Asn Glu Val Asp Ala Leu Ile Ala His
    65                  70                  75                  80
20

    CTG CCA GAA GAC GAG AGG CAG GTC TTG AAG ACG CAG CTG GAG GAG ATC       288
    Leu Pro Glu Asp Glu Arg Gln Val Leu Lys Thr Gln Leu Glu Glu Ile
                        85                  90                  95

25  AAA GTA AAC ATC TCC TTC TTC GGT CTC TGG CGG TAT GCA ACA AAG ATG       336
    Lys Val Asn Ile Ser Phe Phe Gly Leu Trp Arg Tyr Ala Thr Lys Met
                    100                 105                 110

    GAT ATA CTT ATC ATG GTA ATC AGT ACA ATC TGT GCC ATT GCT GCC GCG       384
30  Asp Ile Leu Ile Met Val Ile Ser Thr Ile Cys Ala Ile Ala Ala Ala
                    115                 120                 125

    TCG ACT TTC CAG AGG ATA ATG TTA TAT CAA ATC TCG TAC GAC GAG TTC       432
35  Ser Thr Phe Gln Arg Ile Met Leu Tyr Gln Ile Ser Tyr Asp Glu Phe
            130                 135                 140

    TAT GAT GAA TTG ACC AAG AAC GTA CTG TAC TTC GTA TAC CTC GGT ATC       480
    Tyr Asp Glu Leu Thr Lys Asn Val Leu Tyr Phe Val Tyr Leu Gly Ile
    145                 150                 155                 160
40

    GGC GAG TTT GTC ACT GTC TAT GTT AGT ACT GTT GGC TTC ATC TAT ACC       528
    Gly Glu Phe Val Thr Val Tyr Val Ser Thr Val Gly Phe Ile Tyr Thr
                    165                 170                 175

45  GGA GAA CAC GCC ACG CAG AAG ATC CGC GAG TAT TAC CTT GAG TCT ATC       576
    Gly Glu His Ala Thr Gln Lys Ile Arg Glu Tyr Tyr Leu Glu Ser Ile
                    180                 185                 190

    CTG CGC CAG AAC ATT GGC TAT TTT GAT AAA CTC GGT GCC GGG GAA GTG       624
50  Leu Arg Gln Asn Ile Gly Tyr Phe Asp Lys Leu Gly Ala Gly Glu Val
                    195                 200                 205

    ACC ACC CGT ATA ACA GCC GAT ACA AAC CTT ATC CAG GAT GGC ATT TCG       672
55  Thr Thr Arg Ile Thr Ala Asp Thr Asn Leu Ile Gln Asp Gly Ile Ser
            210                 215                 220

    GAG AAG GTC GGT CTC ACT TTG ACT GCC CTG GCG ACA TTC GTG ACA GCA       720
    Glu Lys Val Gly Leu Thr Leu Thr Ala Leu Ala Thr Phe Val Thr Ala
    225                 230                 235                 240
60

    TTC ATT ATC GCC TAC GTC AAA TAC TGG AAG TTG GCT CTA ATT TGC AGC       768
    Phe Ile Ile Ala Tyr Val Lys Tyr Trp Lys Leu Ala Leu Ile Cys Ser
                    245                 250                 255

65  TCA ACA ATT GTG GCC CTC GTT CTC ACC ATG GGC GGT GGT TCT CAG TTT       816
```

```
              Ser Thr Ile Val Ala Leu Val Leu Thr Met Gly Gly Gly Ser Gln Phe
                      260             265             270

  5   ATC ATC AAG TAC AGC AAA AAG TCG CTT GAC AGC TAC GGT GCA GGC GGC        864
      Ile Ile Lys Tyr Ser Lys Lys Ser Leu Asp Ser Tyr Gly Ala Gly Gly
              275             280             285

      ACT GTT GCG GAA GAG GTC ATC AGC TCC ATC AGA AAT GCC ACA GCG TTT        912
 10   Thr Val Ala Glu Glu Val Ile Ser Ser Ile Arg Asn Ala Thr Ala Phe
          290             295             300

      GGC ACC CAA GAC AAG CTT GCG AAG CAG TAT GAG GTC CAC TTA GAC GAA        960
      Gly Thr Gln Asp Lys Leu Ala Lys Gln Tyr Glu Val His Leu Asp Glu
      305             310             315             320
 15
      GCT GAG AAA TGG GGA ACA AAG AAC CAG ATT GTC ATG GGT TTC ATG ATT       1008
      Ala Glu Lys Trp Gly Thr Lys Asn Gln Ile Val Met Gly Phe Met Ile
                      325             330             335

 20   GGC GCC ATG TTT GGC CTT ATG TAC TCG AAC TAC GGT CTT GGC TTC TGG       1056
      Gly Ala Met Phe Gly Leu Met Tyr Ser Asn Tyr Gly Leu Gly Phe Trp
                  340             345             350

      ATG GGT TCT CGT TTC CTG GTA GAT GGT GCA GTC GAT GTG GGT GAT ATT       1104
 25   Met Gly Ser Arg Phe Leu Val Asp Gly Ala Val Asp Val Gly Asp Ile
                  355             360             365

      CTC ACA GTT CTC ATG GCC ATC TTG ATC GGA TCG TTC TCC TTG GGG AAC       1152
 30   Leu Thr Val Leu Met Ala Ile Leu Ile Gly Ser Phe Ser Leu Gly Asn
          370             375             380

      GTT AGT CCA AAT GCT CAA GCA TTT ACA AAC GCT GTG GCC GCG GCC GCA       1200
      Val Ser Pro Asn Ala Gln Ala Phe Thr Asn Ala Val Ala Ala Ala Ala
      385             390             395             400
 35
      AAG ATA TTT GGA ACG ATC GAT CGC CAG TCC CCA TTA GAT CCA TAT TCG       1248
      Lys Ile Phe Gly Thr Ile Asp Arg Gln Ser Pro Leu Asp Pro Tyr Ser
                  405             410             415

 40   AAC GAA GGG AAG ACG CTC GAC CAT TTT GAG GGC CAC ATT GAG TTA CGC       1296
      Asn Glu Gly Lys Thr Leu Asp His Phe Glu Gly His Ile Glu Leu Arg
                  420             425             430

      AAT GTC AAG CAT ATT TAC CCA TCT AGA CCC GAG GTC ACC GTC ATG GAG       1344
 45   Asn Val Lys His Ile Tyr Pro Ser Arg Pro Glu Val Thr Val Met Glu
              435             440             445

      GAT GTT TCT CTG TCA ATG CCC GCT GGA AAA ACA ACC GCT TTA GTC GGC       1392
      Asp Val Ser Leu Ser Met Pro Ala Gly Lys Thr Thr Ala Leu Val Gly
 50       450             455             460

      CCC TCT GGC TCT GGA AAA AGT ACG GTG GTC GGC TTG GTT GAG CGA TTC       1440
      Pro Ser Gly Ser Gly Lys Ser Thr Val Val Gly Leu Val Glu Arg Phe
      465             470             475             480
 55
      TAC ATG CCT GTT CGC GGT ACG GTT TTG CTG GAT GGC CAT GAC ATC AAG       1488
      Tyr Met Pro Val Arg Gly Thr Val Leu Leu Asp Gly His Asp Ile Lys
                  485             490             495

 60   GAC CTC AAT CTC CGC TGG CTT CGC CAA CAG ATC TCT TTG GTT AGC CAG       1536
      Asp Leu Asn Leu Arg Trp Leu Arg Gln Gln Ile Ser Leu Val Ser Gln
              500             505             510

      GAG CCT GTT CTT TTT GGC ACG ACG ATT TAT AAG AAT ATT AGG CAC GGT       1584
 65   Glu Pro Val Leu Phe Gly Thr Thr Ile Tyr Lys Asn Ile Arg His Gly
```

```
                   515                     520                     525

      CTC ATC GGC ACA AAG TAC GAG AAT GAA TCC GAG GAT AAG GTC CGG GAA        1632
      Leu Ile Gly Thr Lys Tyr Glu Asn Glu Ser Glu Asp Lys Val Arg Glu
 5        530                     535                     540

      CTC ATC GAG AAC GCG GCA AAA ATG GCG AAT GCT CAT GAC TTT ATT ACT        1680
      Leu Ile Glu Asn Ala Ala Lys Met Ala Asn Ala His Asp Phe Ile Thr
      545                     550                     555                 560
10

      GCC TTG CCT GAA GGT TAT GAG ACC AAT GTT GGG CAG CGT GGC TTT CTC        1728
      Ala Leu Pro Glu Gly Tyr Glu Thr Asn Val Gly Gln Arg Gly Phe Leu
                          565                     570                 575

15    CTT TCA GGT GGC CAG AAA CAG CGC ATT GCA ATC GCC CGT GCC GTT GTT        1776
      Leu Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Val Val
                      580                     585                 590

      AGT GAC CCA AAA ATC CTG CTC CTG GAT GAA GCT ACT TCG GCC TTG GAC        1824
20    Ser Asp Pro Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp
                      595                     600                 605

      ACA AAA TCC GAA GGC GTG GTT CAA GCA GCT TTG GAG AGG GCA GCT GAA        1872
      Thr Lys Ser Glu Gly Val Val Gln Ala Ala Leu Glu Arg Ala Ala Glu
25        610                     615                     620

      GGC CGA ACT ACT ATT GTG ATC GCT CAT CGC CTT TCC ACG ATC AAA ACG        1920
      Gly Arg Thr Thr Ile Val Ile Ala His Arg Leu Ser Thr Ile Lys Thr
      625                     630                     635                 640
30

      GCG CAC AAC ATT GTG GTT CTG GTC AAT GGC AAA ATT GCT GAA CAA GGA        1968
      Ala His Asn Ile Val Val Leu Val Asn Gly Lys Ile Ala Glu Gln Gly
                          645                     650                 655

35    ACT CAC GAT GAA TTG GTT GAC CGC GGA GGC GCT TAT CGC AAA CTT GTG        2016
      Thr His Asp Glu Leu Val Asp Arg Gly Gly Ala Tyr Arg Lys Leu Val
                      660                     665                 670

      GAG GCT CAA CGT ATC AAT GAA CAG AAG GAA GCT GAC GCC TTG GAG GAC        2064
40    Glu Ala Gln Arg Ile Asn Glu Gln Lys Glu Ala Asp Ala Leu Glu Asp
                      675                     680                 685

      GCC GAC GCT GAG GAT CTC ACG AAT GCA GAT ATT GCC AAA ATC AAA ACT        2112
      Ala Asp Ala Glu Asp Leu Thr Asn Ala Asp Ile Ala Lys Ile Lys Thr
45        690                     695                     700

      GCG TCA AGC GCA TCA TCC GAT CTC GAC GGA AAA CCC ACA ACC ATT GAC        2160
      Ala Ser Ser Ala Ser Ser Asp Leu Asp Gly Lys Pro Thr Thr Ile Asp
      705                     710                     715                 720
50

      CGC ACG GGC ACC CAC AAG TCT GTT TCC AGC GCG ATT CTT TCT AAA AGA        2208
      Arg Thr Gly Thr His Lys Ser Val Ser Ser Ala Ile Leu Ser Lys Arg
                          725                     730                 735

55    CCC CCC GAA ACA ACT CCG AAA TAC TCA TTA TGG ACG CTG CTC AAA TTT        2256
      Pro Pro Glu Thr Thr Pro Lys Tyr Ser Leu Trp Thr Leu Leu Lys Phe
                      740                     745                 750

      GTT GCT TCC TTC AAC CGC CCT GAA ATC CCG TAC ATG CTC ATC GGT CTT        2304
60    Val Ala Ser Phe Asn Arg Pro Glu Ile Pro Tyr Met Leu Ile Gly Leu
                      755                     760                 765

      GTC TTC TCA GTG TTA GCT GGT GGT GGC CAA CCC ACG CAA GCA GTG CTA        2352
      Val Phe Ser Val Leu Ala Gly Gly Gly Gln Pro Thr Gln Ala Val Leu
65        770                     775                     780
```

```
        TAT GCT AAA GCC ATC AGC ACA CTC TCG CTC CCA GAA TCA CAA TAT AGC        2400
        Tyr Ala Lys Ala Ile Ser Thr Leu Ser Leu Pro Glu Ser Gln Tyr Ser
        785             790             795                 800

  5

        AAG CTT CGA CAT GAT GCG GAT TTC TGG TCA TTG ATG TTC TTC GTG GTT        2448
        Lys Leu Arg His Asp Ala Asp Phe Trp Ser Leu Met Phe Phe Val Val
                    805             810             815

 10     GGT ATC ATT CAG TTT ATC ACG CAG TCA ACC AAT GGT GCT GCA TTT GCC        2496
        Gly Ile Ile Gln Phe Ile Thr Gln Ser Thr Asn Gly Ala Ala Phe Ala
                    820             825             830

        GTA TGC TCC GAG AGA CTT ATT CGT CGC GCG AGA AGC ACT GCC TTT CGG        2544
 15     Val Cys Ser Glu Arg Leu Ile Arg Arg Ala Arg Ser Thr Ala Phe Arg
                    835             840             845

        ACG ATA CTC CGT CAA GAC ATT GCT TTC TTT GAC AAG GAA GAG AAT AGC        2592
        Thr Ile Leu Arg Gln Asp Ile Ala Phe Phe Asp Lys Glu Glu Asn Ser
 20         850             855             860

        ACC GGC GCT CTG ACC TCT TTC CTG TCC ACC GAG ACG AAG CAT CTC TCC        2640
        Thr Gly Ala Leu Thr Ser Phe Leu Ser Thr Glu Thr Lys His Leu Ser
        865             870             875             880

 25
        GGT GTT AGC GGT GTG ACT CTA GGC ACG ATC TTG ATG ACC TCC ACG ACC        2688
        Gly Val Ser Gly Val Thr Leu Gly Thr Ile Leu Met Thr Ser Thr Thr
                    885             890             895

 30     CTA GGA GCG GCT ATC ATT ATT GCC CTG GCG ATT GGG TGG AAA TTG GCC        2736
        Leu Gly Ala Ala Ile Ile Ile Ala Leu Ala Ile Gly Trp Lys Leu Ala
                    900             905             910

        TTA GTT TGT ATC TCG GTT GTG CCG GTT CTC CTG GCA TGC GGT TTC TAC        2784
 35     Leu Val Cys Ile Ser Val Val Pro Val Leu Leu Ala Cys Gly Phe Tyr
                    915             920             925

        CGA TTC TAT ATG CTA GCC CAG TTT CAA TCA CGC TCC AAG CTT GCT TAT        2832
        Arg Phe Tyr Met Leu Ala Gln Phe Gln Ser Arg Ser Lys Leu Ala Tyr
 40         930             935             940

        GAG GGA TCT GCA AAC TTT GCT TGC GAG GCT ACA TCG TCT ATC CGC ACA        2880
        Glu Gly Ser Ala Asn Phe Ala Cys Glu Ala Thr Ser Ser Ile Arg Thr
        945             950             955             960

 45
        GTT GCG TCA TTA ACC CGG GAA AGG GAT GTC TGG GAG ATT TAC CAT GCC        2928
        Val Ala Ser Leu Thr Arg Glu Arg Asp Val Trp Glu Ile Tyr His Ala
                    965             970             975

 50     CAG CTT GAC GCA CAA GGC AGG ACC AGT CTA ATC TCT GTC TTG AGG TCA        2976
        Gln Leu Asp Ala Gln Gly Arg Thr Ser Leu Ile Ser Val Leu Arg Ser
                    980             985             990

        TCC CTG TTA TAT GCG TCG TCG CAG GCA CTT GTT TTC TTC TGC GTT GCG        3024
 55     Ser Leu Leu Tyr Ala Ser Ser Gln Ala Leu Val Phe Phe Cys Val Ala
                    995             1000            1005

        CTC GGG TTT TGG TAC GGA GGG ACA CTT CTT GGT CAC CAC GAG TAT GAC        3072
        Leu Gly Phe Trp Tyr Gly Gly Thr Leu Leu Gly His His Glu Tyr Asp
 60         1010            1015            1020

        ATT TTC CGC TTC TTT GTT TGT TTC TCC GAG ATT CTC TTT GGT GCT CAA        3120
        Ile Phe Arg Phe Phe Val Cys Phe Ser Glu Ile Leu Phe Gly Ala Gln
        1025            1030            1035            1040
 65
```

```
      TCC GCG GGC ACC GTC TTT TCC TTT GCA CCA GAC ATG GGC AAG GCG AAG        3168
      Ser Ala Gly Thr Val Phe Ser Phe Ala Pro Asp Met Gly Lys Ala Lys
                      1045                1050                1055

  5   AAT GCG GCC GCC GAA TTC CGA CGA CTG TTC GAC CGA AAG CCA CAA ATT        3216
      Asn Ala Ala Ala Glu Phe Arg Arg Leu Phe Asp Arg Lys Pro Gln Ile
                      1060                1065                1070

      GAT AAC TGG TCT GAA GAG GGC GAG AAG CTC GAA ACG GTG GAA GGT GAA        3264
 10   Asp Asn Trp Ser Glu Glu Gly Glu Lys Leu Glu Thr Val Glu Gly Glu
                  1075                1080                1085

      ATC GAA TTT AGG AAC GTG CAC TTC AGA TAC CCG ACC CGC CCA GAA CAG        3312
      Ile Glu Phe Arg Asn Val His Phe Arg Tyr Pro Thr Arg Pro Glu Gln
 15       1090                1095                1100

      CCT GTC CTG CGC GGC TTG GAC CTG ACC GTG AAG CCT GGA CAA TAT GTT        3360
      Pro Val Leu Arg Gly Leu Asp Leu Thr Val Lys Pro Gly Gln Tyr Val
      1105                1110                1115                1120
 20
      GCG CTT GTC GGA CCC AGC GGT TGT GGC AAG AGT ACC ACC ATT GCA TTG        3408
      Ala Leu Val Gly Pro Ser Gly Cys Gly Lys Ser Thr Thr Ile Ala Leu
                      1125                1130                1135

 25   CTT GAG CGC TTT TAC GAT GCG ATT GCC GGG TCC ATC CTT GTT GAT GGG        3456
      Leu Glu Arg Phe Tyr Asp Ala Ile Ala Gly Ser Ile Leu Val Asp Gly
                  1140                1145                1150

      AAG GAC ATA AGT AAA CTA AAT ATC AAC TCC TAC CGC AGC TTT CTG TCA        3504
 30   Lys Asp Ile Ser Lys Leu Asn Ile Asn Ser Tyr Arg Ser Phe Leu Ser
                  1155                1160                1165

      CTG GTC AGC CAG GAG CCG ACA CTG TAC CAG GGC ACC ATC AAG GAA AAC        3552
      Leu Val Ser Gln Glu Pro Thr Leu Tyr Gln Gly Thr Ile Lys Glu Asn
 35       1170                1175                1180

      ATC TTA CTT GGT ATT GTC GAA GAT GAC GTA CCG GAA GAA TTC TTG ATT        3600
      Ile Leu Leu Gly Ile Val Glu Asp Asp Val Pro Glu Glu Phe Leu Ile
      1185                1190                1195                1200
 40
      AAG GCT TGC AAG GAC GCT AAT ATC TAC GAC TTC ATC ATG TCG CTC CCG        3648
      Lys Ala Cys Lys Asp Ala Asn Ile Tyr Asp Phe Ile Met Ser Leu Pro
                      1205                1210                1215

 45   GAG GGC TTT AAT ACA GTT GTT GGC AGC AAG GGA GGC ATG TTG TCT GGC        3696
      Glu Gly Phe Asn Thr Val Val Gly Ser Lys Gly Gly Met Leu Ser Gly
                  1220                1225                1230

      GGC CAA AAG CAA CGT GTG GCC ATT GCC CGA GCC CTT CTT CGG GAT CCC        3744
 50   Gly Gln Lys Gln Arg Val Ala Ile Ala Arg Ala Leu Leu Arg Asp Pro
                  1235                1240                1245

      AAA ATC CTT CTT CTC GAT GAA GCG ACG TCA GCC CTC GAC TCC GAG TCA        3792
      Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp Ser Glu Ser
 55       1250                1255                1260

      GAA AAG GTC GTC CAG GCG GCT TTG GAT GCC GCT GCC CGA GGC CGA ACC        3840
      Glu Lys Val Val Gln Ala Ala Leu Asp Ala Ala Ala Arg Gly Arg Thr
      1265                1270                1275                1280
 60
      ACA ATC GCC GTT GCA CAC CGA CTC AGC ACG ATT CAA AAG GCG GAC GTT        3888
      Thr Ile Ala Val Ala His Arg Leu Ser Thr Ile Gln Lys Ala Asp Val
                      1285                1290                1295

 65   ATC TAT GTT TTC GAC CAA GGC AAG ATC GTC GAA AGC GGA ACG CAC AGC        3936
```

```
        Ile Tyr Val Phe Asp Gln Gly Lys Ile Val Glu Ser Gly Thr His Ser
                    1300                1305                1310

  5     GAA CTG GTC CAG AAA AAG GGC CGG TAC TAC GAG CTG GTC AAC TTG CAG         3984
        Glu Leu Val Gln Lys Lys Gly Arg Tyr Tyr Glu Leu Val Asn Leu Gln
                    1315                1320                1325

        AGC TTG GGC AAG GGC CAT                                               4002
        Ser Leu Gly Lys Gly His
 10             1330
```

(2) INFORMATION FOR SEQ ID NO:2:

 15   (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1334 amino acids
                (B) TYPE: amino acid
                (D) TOPOLOGY: linear

 20   (ii) MOLECULE TYPE: protein

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
        Met Ser Pro Leu Glu Thr Asn Pro Leu Ser Pro Glu Thr Ala Met Arg
 25       1               5                   10                  15

        Glu Pro Ala Glu Thr Ser Thr Thr Glu Glu Gln Ala Ser Thr Pro His
                     20                  25                  30

 30     Ala Ala Asp Glu Lys Lys Ile Leu Ser Asp Leu Ser Ala Pro Ser Ser
                     35                  40                  45

        Thr Thr Ala Thr Pro Ala Asp Lys Glu His Arg Pro Lys Ser Ser Ser
           50                  55                  60
 35
        Ser Asn Asn Ala Val Ser Val Asn Glu Val Asp Ala Leu Ile Ala His
          65                  70                  75                  80

        Leu Pro Glu Asp Glu Arg Gln Val Leu Lys Thr Gln Leu Glu Glu Ile
 40                      85                  90                  95

        Lys Val Asn Ile Ser Phe Phe Gly Leu Trp Arg Tyr Ala Thr Lys Met
                    100                 105                 110

 45     Asp Ile Leu Ile Met Val Ile Ser Thr Ile Cys Ala Ile Ala Ala Ala
                    115                 120                 125

        Ser Thr Phe Gln Arg Ile Met Leu Tyr Gln Ile Ser Tyr Asp Glu Phe
            130                 135                 140
 50
        Tyr Asp Glu Leu Thr Lys Asn Val Leu Tyr Phe Val Tyr Leu Gly Ile
        145                 150                 155                 160

        Gly Glu Phe Val Thr Val Tyr Val Ser Thr Val Gly Phe Ile Tyr Thr
 55                     165                 170                 175

        Gly Glu His Ala Thr Gln Lys Ile Arg Glu Tyr Tyr Leu Glu Ser Ile
                    180                 185                 190

 60     Leu Arg Gln Asn Ile Gly Tyr Phe Asp Lys Leu Gly Ala Gly Glu Val
                    195                 200                 205

        Thr Thr Arg Ile Thr Ala Asp Thr Asn Leu Ile Gln Asp Gly Ile Ser
            210                 215                 220
 65
```

Glu Lys Val Gly Leu Thr Leu Thr Ala Leu Ala Thr Phe Val Thr Ala
225   230   235   240

```
5
```
Phe Ile Ile Ala Tyr Val Lys Tyr Trp Lys Leu Ala Leu Ile Cys Ser
                245   250   255

Ser Thr Ile Val Ala Leu Val Leu Thr Met Gly Gly Gly Ser Gln Phe
            260   265   270

```
10
```
Ile Ile Lys Tyr Ser Lys Lys Ser Leu Asp Ser Tyr Gly Ala Gly Gly
        275   280   285

Thr Val Ala Glu Glu Val Ile Ser Ser Ile Arg Asn Ala Thr Ala Phe
    290   295   300

```
15
```
Gly Thr Gln Asp Lys Leu Ala Lys Gln Tyr Glu Val His Leu Asp Glu
305   310   315   320

```
20
```
Ala Glu Lys Trp Gly Thr Lys Asn Gln Ile Val Met Gly Phe Met Ile
                325   330   335

Gly Ala Met Phe Gly Leu Met Tyr Ser Asn Tyr Gly Leu Gly Phe Trp
            340   345   350

```
25
```
Met Gly Ser Arg Phe Leu Val Asp Gly Ala Val Asp Val Gly Asp Ile
            355   360   365

Leu Thr Val Leu Met Ala Ile Leu Ile Gly Ser Phe Ser Leu Gly Asn
    370   375   380

```
30
```
Val Ser Pro Asn Ala Gln Ala Phe Thr Asn Ala Val Ala Ala Ala Ala
385   390   395   400

```
35
```
Lys Ile Phe Gly Thr Ile Asp Arg Gln Ser Pro Leu Asp Pro Tyr Ser
            405   410   415

Asn Glu Gly Lys Thr Leu Asp His Phe Glu Gly His Ile Glu Leu Arg
            420   425   430

```
40
```
Asn Val Lys His Ile Tyr Pro Ser Arg Pro Glu Val Thr Val Met Glu
            435   440   445

Asp Val Ser Leu Ser Met Pro Ala Gly Lys Thr Thr Ala Leu Val Gly
    450   455   460

```
45
```
Pro Ser Gly Ser Gly Lys Ser Thr Val Val Gly Leu Val Glu Arg Phe
465   470   475   480

Tyr Met Pro Val Arg Gly Thr Val Leu Leu Asp Gly His Asp Ile Lys
            485   490   495

```
50
```
Asp Leu Asn Leu Arg Trp Leu Arg Gln Gln Ile Ser Leu Val Ser Gln
            500   505   510

```
55
```
Glu Pro Val Leu Phe Gly Thr Thr Ile Tyr Lys Asn Ile Arg His Gly
            515   520   525

Leu Ile Gly Thr Lys Tyr Glu Asn Glu Ser Glu Asp Lys Val Arg Glu
            530   535   540

```
60
```
Leu Ile Glu Asn Ala Ala Lys Met Ala Asn Ala His Asp Phe Ile Thr
545   550   555   560

Ala Leu Pro Glu Gly Tyr Glu Thr Asn Val Gly Gln Arg Gly Phe Leu
                565   570   575

```
65
```

```
       Leu Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Val Val
                   580             585               590

  5    Ser Asp Pro Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp
                   595             600               605

       Thr Lys Ser Glu Gly Val Val Gln Ala Ala Leu Glu Arg Ala Ala Glu
           610             615               620

 10    Gly Arg Thr Thr Ile Val Ile Ala His Arg Leu Ser Thr Ile Lys Thr
       625             630               635               640

       Ala His Asn Ile Val Val Leu Val Asn Gly Lys Ile Ala Glu Gln Gly
 15                    645               650               655

       Thr His Asp Glu Leu Val Asp Arg Gly Gly Ala Tyr Arg Lys Leu Val
                   660             665               670

 20    Glu Ala Gln Arg Ile Asn Glu Gln Lys Glu Ala Asp Ala Leu Glu Asp
                   675             680               685

       Ala Asp Ala Glu Asp Leu Thr Asn Ala Asp Ile Ala Lys Ile Lys Thr
           690             695               700

 25    Ala Ser Ser Ala Ser Ser Asp Leu Asp Gly Lys Pro Thr Thr Ile Asp
       705             710               715               720

       Arg Thr Gly Thr His Lys Ser Val Ser Ser Ala Ile Leu Ser Lys Arg
 30                    725               730               735

       Pro Pro Glu Thr Thr Pro Lys Tyr Ser Leu Trp Thr Leu Leu Lys Phe
                   740             745               750

 35    Val Ala Ser Phe Asn Arg Pro Glu Ile Pro Tyr Met Leu Ile Gly Leu
                   755             760               765

       Val Phe Ser Val Leu Ala Gly Gly Gly Gln Pro Thr Gln Ala Val Leu
           770             775               780

 40    Tyr Ala Lys Ala Ile Ser Thr Leu Ser Leu Pro Glu Ser Gln Tyr Ser
       785             790               795               800

       Lys Leu Arg His Asp Ala Asp Phe Trp Ser Leu Met Phe Phe Val Val
 45                    805             810               815

       Gly Ile Ile Gln Phe Ile Thr Gln Ser Thr Asn Gly Ala Ala Phe Ala
                   820             825               830

 50    Val Cys Ser Glu Arg Leu Ile Arg Arg Ala Arg Ser Thr Ala Phe Arg
                   835             840               845

       Thr Ile Leu Arg Gln Asp Ile Ala Phe Phe Asp Lys Glu Glu Asn Ser
           850             855               860

 55    Thr Gly Ala Leu Thr Ser Phe Leu Ser Thr Glu Thr Lys His Leu Ser
       865             870               875               880

       Gly Val Ser Gly Val Thr Leu Gly Thr Ile Leu Met Thr Ser Thr Thr
 60                    885               890               895

       Leu Gly Ala Ala Ile Ile Ile Ala Leu Ala Ile Gly Trp Lys Leu Ala
                   900             905               910

 65    Leu Val Cys Ile Ser Val Val Pro Val Leu Leu Ala Cys Gly Phe Tyr
```

```
                915                    920                    925
       Arg Phe Tyr Met Leu Ala Gln Phe Gln Ser Arg Ser Lys Leu Ala Tyr
           930                    935          .          940
   5
       Glu Gly Ser Ala Asn Phe Ala Cys Glu Ala Thr Ser Ser Ile Arg Thr
       945                    950                    955                960
  10   Val Ala Ser Leu Thr Arg Glu Arg Asp Val Trp Glu Ile Tyr His Ala
                       965                    970                    975
       Gln Leu Asp Ala Gln Gly Arg Thr Ser Leu Ile Ser Val Leu Arg Ser
                       980                    985                    990
  15   Ser Leu Leu Tyr Ala Ser Ser Gln Ala Leu Val Phe Phe Cys Val Ala
                       995                    1000                   1005
       Leu Gly Phe Trp Tyr Gly Gly Thr Leu Leu Gly His His Glu Tyr Asp
                   1010                   1015                   1020
  20
       Ile Phe Arg Phe Phe Val Cys Phe Ser Glu Ile Leu Phe Gly Ala Gln
       1025                   1030                   1035                   1040
       Ser Ala Gly Thr Val Phe Ser Phe Ala Pro Asp Met Gly Lys Ala Lys
  25                   1045         ·              1050              ·1055
       Asn Ala Ala Ala Glu Phe Arg Arg Leu Phe Asp Arg Lys Pro Gln Ile
                   1060                   1065                   1070
  30   Asp Asn Trp Ser Glu Glu Gly Glu Lys Leu Glu Thr Val Glu Gly Glu
                   1075                   1080                   1085
       Ile Glu Phe Arg Asn Val His Phe Arg Tyr Pro Thr Arg Pro Glu Gln
           1090                   1095                   1100          .
  35   Pro Val Leu Arg Gly Leu Asp Leu Thr Val Lys Pro Gly Gln Tyr Val
       1105                   1110                   1115                   1120
       Ala Leu Val Gly Pro Ser Gly Cys Gly Lys Ser Thr Thr Ile Ala Leu
                   1125                   1130   ·               1135
  40
       Leu Glu Arg Phe Tyr Asp Ala Ile Ala Gly Ser Ile Leu Val Asp Gly
                   1140                   1145                   1150
  45   Lys Asp Ile Ser Lys Leu Asn Ile Asn Ser Tyr Arg Ser Phe Leu Ser
                   1155                   1160                   1165
       Leu Val Ser Gln Glu Pro Thr Leu Tyr Gln Gly Thr Ile Lys Glu Asn
                   1170                   1175                   1180
  50
       Ile Leu Leu Gly Ile Val Glu Asp Asp Val Pro Glu Glu Phe Leu Ile
       1185                   1190                   1195                   1200
       Lys Ala Cys Lys Asp Ala Asn Ile Tyr Asp Phe Ile Met Ser Leu Pro
  55                   1205                   1210                   1215
       Glu Gly Phe Asn Thr Val Val Gly Ser Lys Gly Gly Met Leu Ser Gly .
                   1220                   1225                   1230
  60   Gly Gln Lys Gln Arg Val Ala Ile Ala Arg Ala Leu Leu Arg Asp Pro
                   1235                   1240                   1245
       Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp Ser Glu Ser
                   1250                   1255                   1260
  65
```

26

```
Glu Lys Val Val Gln Ala Ala Leu Asp Ala Ala Ala Arg Gly Arg Thr
1265             1270             1275             1280
```

5
```
Thr Ile Ala Val Ala His Arg Leu Ser Thr Ile Gln Lys Ala Asp Val
                1285             1290             1295
```

```
Ile Tyr Val Phe Asp Gln Gly Lys Ile Val Glu Ser Gly Thr His Ser
            1300             1305             1310
```

10
```
Glu Leu Val Gln Lys Lys Gly Arg Tyr Tyr Glu Leu Val Asn Leu Gln
        1315             1320             1325
```

```
Ser Leu Gly Lys Gly His
    1330
```

15
(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4002 base pairs

20
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: mRNA

25
    (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: NO

30

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

35
```
AUGUCCCCGC UAGAGACAAA UCCCCUUUCG CCAGAGACUG CUAUGCGCGA ACCUGCUGAG      60

ACUUCAACGA CGGAGGAGCA AGCUUCUACA CCACACGCUG CGGACGAGAA GAAAAUCCUC     120

AGCGACCUCU CGGCUCCAUC UAGUACUACA GCAACCCCCG CAGACAAGGA GCACCGUCCU     180
```

40
```
AAAUCGUCGU CCAGCAAUAA UGCGGUCUCG GUCAACGAAG UCGAUGCGCU UAUUGCGCAC     240

CUGCCAGAAG ACGAGAGGCA GGUCUUGAAG ACGCAGCUGG AGGAGAUCAA AGUAAACAUC     300
```

45
```
UCCUUCUUCG GUCUCUGGCG GUAUGCAACA AAGAUGGAUA UACUUAUCAU GGUAAUCAGU     360

ACAAUCUGUG CCAUUGCUGC CGCGUCGACU UUCCAGAGGA UAAUGUUAUA UCAAAUCUCG     420

UACGACGAGU UCUAUGAUGA AUUGACCAAG AACGUACUGU ACUUCGUAUA CCUCGGUAUC     480
```

50
```
GGCGAGUUUG UCACUGUCUA UGUUAGUACU GUUGGCUUCA UCUAUACCGG AGAACACGCC     540

ACGCAGAAGA UCCGCGAGUA UUACCUUGAG UCUAUCCUGC GCCAGAACAU UGGCUAUUUU     600
```

55
```
GAUAAACUCG GUGCCGGGGA AGUGACCACC CGUAUAACAG CCGAUACAAA CCUUAUCCAG     660

GAUGGCAUUU CGGAGAAGGU CGGUCUCACU UUGACUGCCC UGGCGACAUU CGUGACAGCA     720

UUCAUUAUCG CCUACGUCAA AUACUGGAAG UUGGCUCUAA UUUGCAGCUC AACAAUUGUG     780
```

60
```
GCCCUCGUUC UCACCAUGGG CGGUGGUUCU CAGUUUAUCA UCAAGUACAG CAAAAAGUCG     840

CUUGACAGCU ACGGUGCAGG CGGCACUGUU GCGGAAGAGG UCAUCAGCUC CAUCAGAAAU     900
```

65
```
GCCACAGCGU UUGGCACCCA AGACAAGCUU GCGAAGCAGU AUGAGGUCCA CUUAGACGAA     960
```

```
GCUGAGAAAU GGGGAACAAA GAACCAGAUU GUCAUGGGUU UCAUGAUUGG CGCCAUGUUU    1020
GGCCUUAUGU ACUCGAACUA CGGUCUUGGC UUCUGGAUGG GUUCUCGUUU CCUGGUAGAU    1080
GGUGCAGUCG AUGUGGGUGA UAUUCUCACA GUUCUCAUGG CCAUCUUGAU CGGAUCGUUC    1140
UCCUUGGGGA ACGUUAGUCC AAAUGCUCAA GCAUUUACAA ACGCUGUGGC CGCGGCCGCA    1200
AAGAUAUUUG GAACGAUCGA UCGCCAGUCC CCAUUAGAUC CAUAUUCGAA CGAAGGGAAG    1260
ACGCUCGACC AUUUUGAGGG CCACAUUGAG UUACGCAAUG UCAAGCAUAU UUACCCAUCU    1320
AGACCCGAGG UCACCGUCAU GGAGGAUGUU UCUCUGUCAA UGCCCGCUGG AAAAACAACC    1380
GCUUUAGUCG GCCCCUCUGG CUCUGGAAAA AGUACGGUGG UCGGCUUGGU UGAGCGAUUC    1440
UACAUGCCUG UUCGCGGUAC GGUUUUGCUG GAUGGCCAUG ACAUCAAGGA CCUCAAUCUC    1500
CGCUGGCUUC GCCAACAGAU CUCUUUGGUU AGCCAGGAGC CUGUUCUUUU UGGCACGACG    1560
AUUUAUAAGA AUAUUAGGCA CGGUCUCAUC GGCACAAAGU ACGAGAAUGA AUCCGAGGAU    1620
AAGGUCCGGG AACUCAUCGA GAACGCGGCA AAAAUGGCGA AUGCUCAUGA CUUUAUUACU    1680
GCCUUGCCUG AAGGUUAUGA GACCAAUGUU GGGCAGCGUG GCUUUCUCCU UUCAGGUGGC    1740
CAGAAACAGC GCAUUGCAAU CGCCCGUGCC GUUGUUAGUG ACCCAAAAAU CCUGCUCCUG    1800
GAUGAAGCUA CUUCGGCCUU GGACACAAAA UCCGAAGGCG UGGUUCAAGC AGCUUUGGAG    1860
AGGGCAGCUG AAGGCCGAAC UACUAUUGUG AUCGCUCAUC GCCUUUCCAC GAUCAAAACG    1920
GCGCACAACA UUGUGGUUCU GGUCAAUGGC AAAAUUGCUG AACAAGGAAC UCACGAUGAA    1980
UUGGUUGACC GCGGAGGCGC UUAUCGCAAA CUUGUGGAGG CUCAACGUAU CAAUGAACAG    2040
AAGGAAGCUG ACGCCUUGGA GGACGCCGAC GCUGAGGAUC UCACGAAUGC AGAUAUUGCC    2100
AAAAUCAAAA CUGCGUCAAG CGCAUCAUCC GAUCUCGACG GAAAACCCAC AACCAUUGAC    2160
CGCACGGGCA CCCACAAGUC UGUUUCCAGC GCGAUUCUUU CUAAAGACC CCCCGAAACA     2220
ACUCCGAAAU ACUCAUUAUG GACGCUGCUC AAAUUUGUUG CUUCCUUCAA CCGCCCUGAA    2280
AUCCCGUACA UGCUCAUCGG UCUUGUCUUC UCAGUGUUAG CUGGUGGUGG CCAACCCACG    2340
CAAGCAGUGC UAUAUGCUAA AGCCAUCAGC ACACUCUCGC UCCCAGAAUC ACAAUAUAGC    2400
AAGCUUCGAC AUGAUGCGGA UUUCUGGUCA UUGAUGUUCU CGUGGUUGG UAUCAUUCAG     2460
UUUAUCACGC AGUCAACCAA UGGUGCUGCA UUUGCCGUAU GCUCCGAGAG ACUUAUUCGU    2520
CGCGCGAGAA GCACUGCCUU UCGGACGAUA CUCCGUCAAG ACAUUGCUUU CUUUGACAAG    2580
GAAGAGAAUA GCACCGGCGC UCUGACCUCU UUCCUGUCCA CCGAGACGAA GCAUCUCUCC    2640
GGUGUUAGCG GUGUGACUCU AGGCACGAUC UUGAUGACCU CCACGACCCU AGGAGCGGCU    2700
AUCAUUAUUG CCCUGGCGAU UGGGUGGAAA UUGGCCUUAG UUUGUAUCUC GGUUGUGCCG    2760
GUUCUCCUGG CAUGCGGUUU CUACCGAUUC UAUAUGCUAG CCCAGUUUCA AUCACGCUCC    2820
AAGCUUGCUU AUGAGGGAUC UGCAAACUUU GCUUGCGAGG CUACAUCGUC UAUCCGCACA    2880
```

```
      GUUGCGUCAU UAACCCGGGA AAGGGAUGUC UGGGAGAUUU ACCAUGCCCA GCUUGACGCA    2940

      CAAGGCAGGA CCAGUCUAAU CUCUGUCUUG AGGUCAUCCC UGUUAUAUGC GUCGUCGCAG    3000

 5    GCACUUGUUU UCUUCUGCGU UGCGCUCGGG UUUUGGUACG GAGGGACACU UCUUGGUCAC    3060

      CACGAGUAUG ACAUUUUCCG CUUCUUUGUU UGUUUCUCCG AGAUUCUCUU UGGUGCUCAA    3120

      UCCGCGGGCA CCGUCUUUUC CUUUGCACCA GACAUGGGCA AGGCGAAGAA UGCGGCCGCC    3180

10    GAAUUCCGAC GACUGUUCGA CCGAAAGCCA CAAAUUGAUA ACUGGUCUGA AGAGGGCGAG    3240

      AAGCUCGAAA CGGUGGAAGG UGAAAUCGAA UUUAGGAACG UGCACUUCAG AUACCCGACC    3300

15    CGCCCAGAAC AGCCUGUCCU GCGCGGCUUG GACCUGACCG UGAAGCCUGG ACAAUAUGUU    3360

      GCGCUUGUCG ACCCAGCGG UUGUGGCAAG AGUACCACCA UUGCAUUGCU UGAGCGCUUU    3420

      UACGAUGCGA UUGCCGGGUC CAUCCUUGUU GAUGGGAAGG ACAUAAGUAA ACUAAAUAUC    3480

20    AACUCCUACC GCAGCUUUCU GUCACUGGUC AGCCAGGAGC CGACACUGUA CCAGGGCACC    3540

      AUCAAGGAAA ACAUCUUACU UGGUAUUGUC GAAGAUGACG UACCGGAAGA AUUCUUGAUU    3600

25    AAGGCUUGCA AGGACGCUAA UAUCUACGAC UUCAUCAUGU CGCUCCCGGA GGGCUUUAAU    3660

      ACAGUUGUUG GCAGCAAGGG AGGCAUGUUG UCUGGCGGCC AAAAGCAACG UGUGGCCAUU    3720

      GCCCGAGCCC UUCUUCGGGA UCCCAAAAUC CUUCUUCUCG AUGAAGCGAC GUCAGCCCUC    3780

30    GACUCCGAGU CAGAAAAGGU CGUCCAGGCG GCUUUGGAUG CCGCUGCCCG AGGCCGAACC    3840

      ACAAUCGCCG UUGCACACCG ACUCAGCACG AUUCAAAAGG CGGACGUUAU CUAUGUUUUC    3900

35    GACCAAGGCA AGAUCGUCGA AAGCGGAACG CACAGCGAAC UGGUCCAGAA AAAGGGCCGG    3960

      UACUACGAGC UGGUCAACUU GCAGAGCUUG GGCAAGGGCC AU                       4002
```